



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/829,831	04/10/2001	Elizabeth Shriberg	SRI/4316	1269
7590	04/07/2004		EXAMINER	
Thomason, Moser & Patterson LLP Attorneys At Law First Floor 595 Shrewsbury Avenue Shrewsbury, NJ 07702			LEWIS, MICHAEL A	
			ART UNIT	PAPER NUMBER
			2655	
			DATE MAILED: 04/07/2004	

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/829,831	SHRIBERG ET AL.	
	Examiner	Art Unit	
	Lewis A Michael	2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) Responsive to communication(s) filed on ____.
- 2a) This action is FINAL. 2b) This action is non-final.
- 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) Claim(s) 1-21 is/are pending in the application.
- 4a) Of the above claim(s) ____ is/are withdrawn from consideration.
- 5) Claim(s) ____ is/are allowed.
- 6) Claim(s) 1-21 is/are rejected.
- 7) Claim(s) ____ is/are objected to.
- 8) Claim(s) ____ are subject to restriction and/or election requirement.

Application Papers

- 9) The specification is objected to by the Examiner.
- 10) The drawing(s) filed on ____ is/are: a) accepted or b) objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
 - a) All b) Some * c) None of:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. ____.
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|---|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413) |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date: ____. |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date <u>04</u> . | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
| | 6) <input type="checkbox"/> Other: ____. |

DETAILED ACTION

Claim Rejections - 35 USC § 103

1. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

2. The factual inquiries set forth in *Graham v. John Deere Co.*, 383 U.S. 1, 148 USPQ 459 (1966), that are applied for establishing a background for determining obviousness under 35 U.S.C. 103(a) are summarized as follows:

1. Determining the scope and contents of the prior art.
2. Ascertaining the differences between the prior art and the claims at issue.
3. Resolving the level of ordinary skill in the pertinent art.
4. Considering objective evidence present in the application indicating obviousness or nonobviousness.

3. Claims 1,2,11,12 & 21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Lee et al (U.S. Patent 5617507) in view of Silverman (U.S. Patent 5890117).

Regarding claims 1, 11 & 21, Lee et al. discloses a method, apparatus and computer processor with storage (Fig 1 (10), Col 10, Lines 5 – 10) for processing a speech signal comprising:

a. Extracting prosodic features from a speech signal (Fig. 1(1,2)).

Lee et al. do not explicitly disclose the following:

- a. Modeling the prosodic features to identify at least one speech endpoint
- b. Producing an endpoint signal corresponding to the occurrence of the at least one speech endpoint.

However, Silverman teaches a series of rules [model] that produces prosodically annotated text that has information related to speech boundaries [including claimed endpoints] (Fig. 3 and Col 11, 1 – 35). Prosodic information is used to determine boundary conditions in speech and it is essential for the accurate recognition and synthesis of speech.

Therefore it would have been obvious to one of ordinary skill at the time of the invention to modify Lee et al. to explicitly teach a prosodic modeling as taught by Silverman since it would have improved the quality of the synthesized speech signal produced.

Regarding claim 2 & 12, the modified Lee et al disclose the extracting step that comprises of processing pitch information within the speech signal (Fig. 4).

4. Claims 3, 4,13 & 14 are rejected under 35 U.S.C. 103(a) as being unpatentable over Lee et al (U.S. Patent 5617507) in view of Silverman (U.S. Patent 5890117) and in further view Chihara (U.S. Patent 6470316).

Regarding claim 3 & 13, the modified Lee et al disclose the extracting step further comprises: determining a duration pattern (Fig 7(10)). Lee et al. do not disclose performing pause analysis. However, Chihara teaches a prosody generation module that assesses duration and pause features for phonemes (Col 6, Lines 40 – 45; Fig 2 (206)). Analysis related to duration and pauses are necessary to synthesize natural sounding speech.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the modified Lee to analyze the duration and pause features in speech as taught by Chihara since it would have been beneficial to speech synthesis and recognition applications.

Regarding claim 4 & 14, Lee et al. do not disclose the processing step comprises: generating a pitch contour; producing a pitch movement model from the pitch contour; and extracting at least one pitch parameter from the pitch movement model. However, Chihara teaches a prosody generation module that extracts the pitch contour information [pitch movement model] (Fig 2 (202)) that includes features such as the start point, end point and magnitude of the pitch within an analysis window. Pitch contour information is key to prosody modeling

since the dynamics of features covers emotions, etc. that produces a good prosodic model.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the modified Lee to analyze the pitch contour information as taught by Chihara since it would have produced a more accurate model for speech.

5. Claim 5 & 15 are rejected under 35 U.S.C. 103(a) as being unpatentable over Lee et al (U.S. Patent 5617507) in view of Silverman (U.S. Patent 5890117) in view of Chihara (U.S. Patent 6470316) and in further view Lin (U.S. Patent 4799261).

Regarding claim 5 & 15, the modified Lee et al. do not explicitly disclose that at least one pitch parameter is a pitch movement slope. However, Lin et al. teach the use of extracting the pitch slope [claimed pitch movement slope] from a pitch track (Col 7, Lines 56-68). Information related to pitch contour and intonation helps with the naturalness and intelligibility of encoded speech.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the modified Lee to extract the pitch slope information as taught by Lin et al. since it would have produced a more accurate model for speech.

6. Claims 6 & 16 are rejected under 35 U.S.C. 103(a) as being unpatentable over Lee et al (U.S. Patent 5617507) in view of Silverman (U.S. Patent 5890117) in view of Chihara (U.S. Patent 6470316) and in further view Chihara (U.S. Patent 6625575).

Regarding claim 6 & 16, Lee et al. do not disclose that at least one pitch parameter is a difference between the pitch information in the speech signal and baseline pitch information. However, Chihara teaches the use of the difference of the pitch and base pitch information to modify the intonation process in a text to speech conversion system (Col 19, Line 65 – Col 2, Line 5). Control of the intonation of speech gives the produced speech a more natural and intelligible sound.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the modified Lee to take the difference of the pitch and the baseline pitch information as taught by Lin et al. since it would have produced a more accurate model for speech.

7. Claims 7,8,9,10,17,18,19 & 20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Lee et al (U.S. Patent 5617507) in view of Silverman (U.S. Patent 5890117) and in further view Neumeyer et al (U.S. Patent 6226611).

Regarding claim 7 & 17, Lee et al do not disclose the producing step comprising generating a posterior probability regarding the at least one speech endpoint. However, Neumeyer et al. teach the use of acoustic unit duration scorer that is continuously updated that uses probability to calculate the duration of speech segments and sets the time boundaries [endpoint detection] (Col 5, Line 55 – Col 6, Line5; Col 7, Lines 28 – 45). Endpoint detection is important in preprocessing speech for applications related to speech synthesis and word spotting/recognition applications.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the modified Lee by generating a posterior probability using endpoint information as taught by Neumeyer et al. since it would have produced a more accurate model for speech.

Regarding claims 8 & 18, Lee et al do not disclose that the posterior probability regarding a plurality of speaker states including a probability that a speaker has completed an utterance, a probability that the speaker is pausing due to hesitation, or a probability that the speaker is talking fluently. However, Neumeyer et al. teach an HMM model incorporated with the acoustic duration scorer [posterior probability device] that has the ability to distinguish pauses that occur during words. These models also include the context-dependent features

[which includes hesitation] etc. (Col 9, Line 45 – 50). Generally, the ability to judge speaker state including the context of paused etc. is important for accurate speech recognition systems.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the modified Lee by including a probability function that predict the end of an utterance, pauses or fluency as taught by Neumeyer et al. since it would have produced a more accurate model for speech.

Regarding claim 9 & 19, Lee et al do not disclose continuously generating a posterior probability as the speech is being processed. However, Neumeyer et al. teach the use of acoustic unit duration scorer that is continuously updated that uses probability to calculate the duration of speech segments and sets the time boundaries [endpoint detection] (Col 5, Line 55 – Col 6, Line5; Col 7, Lines 28 – 45). Endpoint detection is important in preprocessing speech in applications related to speech synthesis and word spotting/recognition.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the modified Lee to process speech by continuously updating a posterior probability as taught by Neumeyer et al. since it would have produced a more accurate model for speech.

Regarding claims 10 & 20, Lee et al do not disclose executing a speech recognition routine for processing the speech signal using at least one speech endpoint. However, Neumeyer et al. teach the use of a speech recognizer that utilizes the acoustic unit duration scorer which sets the endpoint (Fig. 3; Col 9, Line 29 - 50). Pronunciation scores based on the duration of acoustic units including end-pointing detection is an important feature in improving speaker independent speech recognition system.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the modified Lee to use a speech recognizer with end-pointing as taught by Neumeyer et al. because proper endpoint prediction would have resulted in a more accurate speech recognizer.

Conclusion

1. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Chihara	U.S. Patent (6470316)
Boss	U.S. Patent (5933805)
Shao et al.	U.S. Patent Application (20020049593)
Bellegarda et al.	U.S. Patent (5121428)
Acero	U.S. Patent (6253182)

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Michael A Lewis whose telephone number is 703 305-8730. The examiner can normally be reached on Monday through Friday, 8:30 am – 5 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, To Doris can be reached on (703)305-4827. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Lewis A Michael
Examiner
Art Unit 2655

Mal

3/16/2004


DORIS H. TO *4/15/04*
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2600